International Journal of Infectious Diseases 46 (2016) 49–55

Contents lists available at ScienceDirect

International Journal of Infectious Diseases

journal homepage: www.elsevier.com/locate/ijid

Association of ecological factors with Rift Valley fever occurrence and mapping of risk zones in Kenya

Gladys Mosomtai [a], Magnus Evander [b], Per Sandström [c], Clas Ahlm [d], Rosemary Sang [a], Osama Ahmed Hassan [b], Hippolyte Affognon [a], Tobias Landmann [a,*]

[a] International Centre of Insect Physiology and Ecology, PO Box 30772-00100, Nairobi, Kenya
[b] Department of Clinical Microbiology, Virology, Umeå University, Umeå, Sweden
[c] Department of Forest Resource Management, Faculty of Forest Sciences, Swedish University of Agricultural Sciences, Umeå, Sweden
[d] Department of Clinical Microbiology, Infectious Diseases, Umeå University, Umeå, Sweden

ARTICLE INFO

Article history:
Received 7 December 2015
Received in revised form 1 February 2016
Accepted 14 March 2016

Corresponding Editor: Eskild Petersen, Aarhus, Denmark.

Keywords:
Rift Valley fever
Evapotranspiration
Normalized difference vegetation index
Animal density
Disease mapping

SUMMARY

Objective: Rift Valley fever (RVF) is a mosquito-borne infection with great impact on animal and human health. The objectives of this study were to identify ecological factors that explain the risk of RVF outbreaks in eastern and central Kenya and to produce a spatially explicit risk map.
Methods: The sensitivity of seven selected ecological variables to RVF occurrence was assessed by generalized linear modelling (GLM). Vegetation seasonality variables (from normalized difference vegetation index (NDVI) data) and 'evapotranspiration' (ET) (metrics) were obtained from 0.25–1 km MODIS satellite data observations; 'livestock density' (N/km^2), 'elevation' (m), and 'soil ratio' (fraction of all significant soil types within a certain county as a function of the total area of that county) were used as covariates.
Results: 'Livestock density', 'small vegetation integral', and the second principal component of ET were the most significant determinants of RVF occurrence in Kenya (all $p \leq 0.01$), with high RVF risk areas identified in the counties of Tana River, Garissa, Isiolo, and Lamu.
Conclusions: Wet soil fluxes measured with ET and vegetation seasonality variables could be used to map RVF risk zones on a sub-regional scale. Future outbreaks could be better managed if relevant RVF variables are integrated into early warning systems.

1. Introduction

Kenya has experienced several outbreaks of Rift Valley fever (RVF), resulting in human disease with a high case fatality[1] and considerable loss of livestock.[2] The disease is caused by the Rift Valley fever virus (RVFV), which is transmitted to vertebrates through the bites of the mosquito vector and through contact with the body fluids of infested animals.[3] In general, RVF outbreaks are triggered by periods of above normal rainfall events and higher temperatures. Outbreaks typically occur at 5–15-year intervals, and the occurrence is sporadic in inter-epidemic/epizootic periods.[4] However, little is known about the role of key ecological determinants of RVF on the landscape and regional scales and regarding the exploration of spatially explicit models for risk mapping.[1,5]

Large scale and data-driven RVF occurrence studies in Africa have relied primarily on modelling approaches that use ecological proxies such as climate data or vegetation activity averages over a certain period (i.e., normalized difference vegetation index (NDVI) data metrics).[2,6] These modelling studies have not, in the most part, explored the use of spatially explicit, localized, and temporally varying ecological factors to assess risk zones.[7,8] Spatially varying proxies for ecological processes on inter-annual vegetation seasonality and 'actual' (as opposed to modelled) land surface fluxes from water bodies would greatly improve RVF occurrence and risk zone modelling (regional scale) and mapping (local to landscape scales). Specifically, RVF occurrence on the landscape scale is driven largely by inter-annual changes in flooding and vegetation density dynamics.[9] Ecological variables are key determinants of mosquito habitat availability.[2,10] Satellite imagery offers the ability to derive 'actual' land surface dynamics information, which improves the mapability of specific ecological variables.[11] Fine spatially and temporally, as well as well-processed remote sensing-based datasets, essentially help

* Corresponding author. Tel.: +254 715 286 949; fax: +254 20 8632001/2.
E-mail address: tlandmann@icipe.org (T. Landmann).

http://dx.doi.org/10.1016/j.ijid.2016.03.013

to reduce model over-fitting (over-predicting), which thus enhances the meaningfulness and accuracy of disease modelling outputs.[12]

In this study, RVF occurrence and risk zones were primarily defined by the ecological conditions that ascribe vector habitat suitability and vector propagation. It is acknowledged that there are specific 'non-mapable' socio-economic and cultural factors and risk dimensions, including meat handling procedures of infected animals, seroprevalence in livestock (number of infected animals), and small-scale herd density and migration patterns.[3,6] RVF-relevant ecological factors have the advantage that most of them can be mapped effectively over larger areas and be used as disease trigger mechanisms in early warning systems.[13]

For RVF occurrence, ecologically driven risks are related to primary and secondary vector habitat availability and dynamics.[14] Above average rainfall events (i.e., El Niño-Southern Oscillation (ENSO) events) usually trigger flooding and enhanced vegetation growth, which enable the breeding and propagation of secondary RVF vectors in particular.[2] Thus, if ecological trigger variables (as proxies) for RVF outbreaks can be recognized and the interactions between these variables (factors) can be identified, disease occurrence or risk zone maps can be produced.[15] Risk maps in disease mapping and modelling refer to the differentiation of endemic- from epidemic-prone and non-epidemic areas in time and space.[16]

Most previous studies on RVF in Kenya have used sea surface temperature (SST) abnormalities, climate variables, coarse resolution (>1 km pixel resolution) NDVI metrics (averages over 1 year), and the presence of hydrographic soils to model RVF occurrence and risk zones.[2,6,17] Most studies have made the assumption that there is a causal relationship between green vegetation development (i.e., NDVI) and vector breeding spaces on a regional scale. However, no attempt has yet been made to include other spatial invariant and more relevant remote sensing variables over larger areas and to investigate the seasonality of NDVI at moderate pixel resolutions (<300 m) in order to better mimic the temporal dynamics of vegetation dynamics. Moreover, there is a need to use intrinsic socio-ecological factors (other than climate), such as livestock densities, in data-driven RVF occurrence mapping approaches.[18]

2. Methods

2.1. Ecological setting and epidemiology of RVF in the study area

The study area stretches from the eastern to the central part of Kenya and covers the newly created counties of Baringo, Laikipia, Meru, Isiolo, Garissa, Tana River, and Lamu (Figure 1), spanning over 142 745 km^2 (Figure 1). Although semi-arid, this region is prone to large-scale flooding during the two rainy seasons of April to May and November to December.[19] The study area is slightly undulating with large tracts of black cotton and alluvial soils that are known to exhibit high water retention potential. Flooding and consequently mosquito hatching often occur in water-filled topographic depressions, or so called 'dambos'.[9] The most dominant natural woody species are acacias, which occur alongside open grasslands; the predominant land use is pastoralism.[20]

The region has been prone to multiple epizootics and epidemics since the 1960s and exhibits seasonal flooding conditions that provide ideal breeding conditions for the primary and secondary RVF vectors.[13] The study region was also selected because some districts had experienced two RVF outbreaks, one in 1997/1998 and one in 2006/2007, while other districts, such as Baringo, were newly affected in the 2006/2007 outbreak period.[21,22]

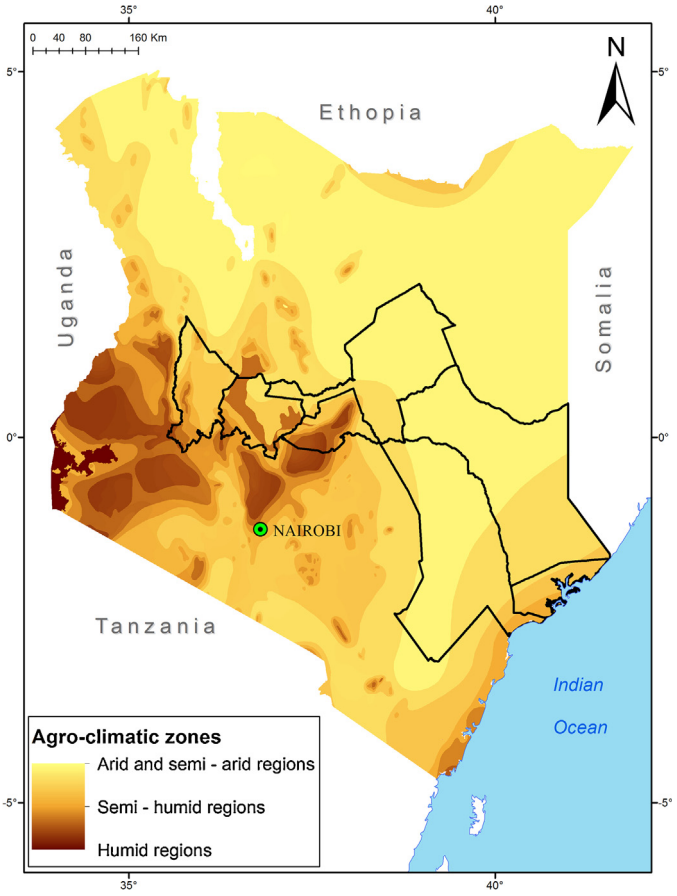

Figure 1. Agro-ecological zone map for Kenya showing the counties, outlined in black, constituting the study region (RVF occurrence area).

2.2. Methodological approach and risk mapping approach

2.2.1. Overview of variables and selection criteria

Table 1 shows the data characteristics of the ecological variables (covariates) used. Each of the covariates is explained in the sections below. The 'soil ratio' was derived from a geographical information system (GIS) vector data layer; otherwise all covariates used in the modelling were derived from pixel-based raster datasets. The two remote sensing variables (evapotranspiration (ET) and NDVI) varied temporally and spatially according to pixel sizes and the observation time frames and frequencies.

The best available datasets were chosen as covariates in terms of spatial resolution, consistency, and temporal alignment with the

Table 1
Ecological covariates (variables) that were used in the statistical analysis

Variable name	Units	Resolution and source
Animal density	Numbers/km^2	5 km (FAO)
Elevation	m (above mean sea level)	90 m SRTM
Season length	N/A	250 m MODIS NDVI
Small integral	N/A	250 m MODIS NDVI
Soil ratio	N/A	5 km (Kenya Soil Survey)[a]
PC1_ET	N/A	1 km MODIS ET
PC2_ET	N/A	1 km MODIS ET

FAO, Food and Agriculture Organization; SRTM, Shuttle Radar Topography Mission; MODIS, moderate resolution imaging spectroradiometer; NDVI, normalized difference vegetation index; ET, evapotranspiration; PC1, first principal component; PC2, second principal component; N/A, non-applicable (unit-less).

[a] Derived from a geographical information systems data layer.

RVF occurrence data used. 'Animal density', 'elevation', and 'soil ratio' and the NDVI-derived variables were chosen because of their known sensitivity to RVF occurrence throughout eastern Africa.[6,9] The remote sensing-based variables (NDVI-derived vegetation integral and length of growing season and ET) were included as 'new' variables, since they mimic the ecological habitat conditions of the mosquito (vector) breeding sites very well. The two NDVI-based seasonality variables were preferred over other NDVI variables (such as maximum NDVI value per season), because they mimic and discriminate abnormal vegetation 'greening' conditions due to high amounts of rainfall in semi-arid regions very well.[23] ET as an ecological variable is overly sensitive to land surface fluxes that can be ascribed to conducive ecological conditions for vector breeding and propagation,[24] such as wet soil conditions or fluxes from inundated areas. Although ecologically meaningful, spatial and temporally variant vegetation seasonality variables and ET have never been tested as independent ecological variables in any disease mapping study in Kenya.

Collinearity between the variables was investigated using per district variable means over the observation period (2001–2013) and linear regression with correlation coefficient thresholds of $|r| > 0.7$.[25] None of the proposed ecological variables showed a statistical collinearity between them using the correlation threshold mentioned.

The covariates were processed to represented mean values for each county in Kenya. For the satellite variables, the county means were representative of the time period in which the satellite data were captured, i.e. from 2001 to 2013. The county means were used for the variable sensitivity assessment and subsequent risk mapping.

2.2.2. Processing of satellite datasets and variables

In this study, satellite-derived time-series datasets of NDVI and ET were processed and used as ecological variables (covariates) (Table 1). Sixteen-day composite images of NDVI from the 250-m MODIS MOD13Q1 product (collection 5) for the years 2001–2013 were pre-processed to reduce residual noise such as clouds and cloud shadow.[26] The small seasonal integral ('small integral') and length of the main growing period ('length of season') were derived from the corrected time-series NDVI data using the TIMESAT tool.[27] The small vegetation integral is the magnitude of the accumulated seasonal vegetation productivity.[23] Both vegetation seasonality variables were derived as means over the 13-year observation period. Eight-day composite ET imagery (best night and day time observation within an 8-day period) from the 1-km MOD16 data product were acquired for 2001 to 2013. No pre-processing was performed for the ET time-series data, since ET is estimated from an array of pre-processed MODIS products and other modelled environmental data variables.[24] Lastly, a principal component analysis was performed on the ET time-series data in order to discern the main data variability over the observation period. Elevated ET values during the 2006/2007 RVF outbreak period should thus be reflected in the main principal components (PC1 and PC2). PC1 and PC2 were selected as covariates for the generalized linear modelling (GLM) (Table 1).

Livestock density data for Kenya (n/km^2; cumulative for goats and sheep, camels, and cattle) were acquired from the Food and Agriculture Organization's gridded livestock of the world dataset.[28] Mean livestock density was computed for each of the counties from the 5-km grid cell livestock density data (expressed as n/km^2 per county).

The 90-m digital elevation model (DEM) data were acquired from the Shuttle Radar Topography Mission (SRTM). The SRTM DEM data were sourced from the United States Geological Survey (USGS) data archive (https://lta.cr.usgs.gov/SRTM1Arc). The 'void filled' DEM data are corrected for missing data values using data interpolation and fill values. Mean elevation per county/district was used instead of DEM-derived topology factors, since it would not have been feasible to derive accurate and comparable county means for a particular topology.

Digital soil type data were obtained from the Kenya Soil Survey dataset.[21] The soil polygon data were converted to raster data. Due to the soil data characteristics (not normally distributed counts data), a negative binomial model was used to analyze the influence of a particular soil type and the area covered by that particular soil type on RVF occurrence within a particular county. RVF occurrence was derived from the historically documented number of outbreaks recorded within certain counties.[13] Soil types with the highest and most statistically significant influences on RVF outbreaks (p-values >0.05) were tagged, i.e. pre-selected. These soil types were alisols, calcisols, greyzems, leptosols, lixisols, luvisols, planosols, solonchaks, and solonetz. The soil ratio (Table 1) is essentially the fraction of all significant soil types within a certain county as a function of the total area of that county.

2.2.3. Generalized linear modelling (GLM) and risk mapping

All 47 counties in Kenya were used for the sensitivity analysis using statistical GLM. A RVF risk map was then derived for the semi-arid study region (Figure 1) using only the most significant (important) variables from GLM. GLM was used since the data were not normally distributed, were empirical in nature, and comprised a mixture of discrete counts and continuous data variables.[29]

Two sets of dependent data variables were used. Firstly, the number of RVF outbreaks a district (county) had recorded between 1951 and 2007[13] (counts data; $n = 46$, range 0–23, mean 9, standard deviation 10.5), and secondly, the presence or absence (binary) of RVF within a certain county during the 2006/2007 outbreak period. GLM was used to link the two response variables to the above-mentioned candidate covariates. Specifically, a negative binomial (maximum likelihood) GLM model was used for the counts data, while a binary logistic regression model was used for the absence and presence binary data.[29] Because of over-dispersion in the count data, a negative binomial model was preferred over a Poisson model. In both GLM models, significance levels (as p-values) were computed for each of the covariates described in section 2.2.2.

The probability levels of each covariate, set at $p < 0.01$ for the covariates in the counts model and at $p < 0.05$ for the variables in the binary model, were determined for the two GLM models. The probability levels could be different because they were derived from two different GLM models and response (dependent) variables. The significances of the binomial model were, moreover, more meaningful, comparable, and stratified when setting higher significance levels. Furthermore, percentage root mean square errors (% RMSE) and the residuals (plots) were computed for both GLM models. The RMSE is a single measure of the predictive power of a model that uses the differences between values predicted by the model and the values actually observed (i.e., the residuals in the model).[30] For model to model comparability reasons, the RMSE was expressed as a percentage. The Akaike information criterion (AIC) scores were used as a secondary sensitivity measure for the ecological variables.

Subsequently, both GLM models were re-computed using only the three most significant covariates. The re-computed models (after variable refinement) were compared to the original models to ascertain the improvement in model fit and model significances. The residuals and the % RMSE scores, as well as the overall model significances (intercept-only p-values), were used for the comparison.

Furthermore, the re-computed models were validated by evaluating the respective regression deviances. The regression deviance is the likelihood ratio between a fully fitted model and

the modelled data from data collections (i.e., the ecological variables). Regression deviances can be used as an intrinsic goodness-of-fit validation measure for a given statistical model.[31]

Before assimilating the most significant variables and classifying risk zones, the raster data were re-classified into five vulnerability classes using the Jenks natural break algorithm,[32] which maximizes the variance in the data for subsequent classification. This was done for the study region. Using a weighted sum tool, the three most significant variables, based on their *p*-values, were weighted and combined.[32] The following weightings were determined: 0.5 for the most significant variable, 0.3 for the second most significant variable, and 0.2 for the third most significant variable.

3. Results

3.1. Significances and relative importance of variables

'Animal density' ($p = 0.004$), 'small integral' ($p = 0.001$), and 'PC2_ET' ($p = 0.008$) were the most significant ecological variables that explained RVF occurrence in the study region when using the negative binomial GLM (Table 2). The results from the binary logistic regression model re-affirmed the significance of 'small integral' ($p < 0.05$; Table 2), while for 'animal density' only, the relatively lower AIC score in the binary GLM (AIC = 43.91; Table 2) confirmed the relative importance of this variable.

The residual plots and the % RMSE for both GLM models (Figures 2 and 3) confirmed the relative importance of the three most significant variables from the negative binomial GLM ('animal density', 'small integral', and 'PC2_ET'). Compared to Figure 2A, Figure 2B shows a lower variance around zero and a lower % RMSE in the negative binomial GLM.

In Figure 2B only the three most significant covariates are used, while in Figure 2A the four insignificant variables (Table 2) are used in the negative binomial model run. Essentially, 66% of the data points were, before variable selection, located within a 5% value buffer area in close proximity to the zero line (Figure 2A), whereas after variable selection and re-modelling, 84% of the data points were within this same buffer area (Figure 2B). Likewise in Figure 3B, the residuals are non-randomly and more closely distributed around the zero line and the lower % RMSE illustrates a more accurate model prediction for the response variables after the aforementioned variables ('animal density', 'small integral', and 'PC2_ET') were used in the binary modelling. The % RMSE exhibited a value of 66 before variable selection (Figure 3A) in comparison with 38 after performing variable selection and re-running the binary GLM (Figure 3B). In both GLM models, the intercept-only *p*-values also decreased when using the three aforementioned most significant variables. For the negative binomial model, for instance, the *p*-value was 0.15 before variable selection and 0.063 after variable selection.

Table 2
Covariates and generalized linear modelling results for all 47 counties using the historical Rift Valley fever counts per district/county data (negative binomial) and the county-based presence/absence of Rift Valley fever during the 2006/2007 outbreak period (binary logistical)

Covariates (variables)	$p < 0.01$		$p < 0.05$	
	Negative binomial		Binary logistical	
	p-Value	AIC	*p*-Value	AIC
Animal density	0.004	290.02	0.19	43.91
Elevation	0.103	292.68	0.79	62.03
Season length	0.011	292.91	0.58	65.29
Small integral	0.001	280.78	0.04	48.10
Soil ratio	0.021	289.40	0.53	62.45
PC1_ET	0.582	732.31	0.30	56.93
PC2_ET	0.008	292.97	0.56	59.54

AIC, Akaike information criterion; PC1, first principal component; PC2, second principal component; ET, evapotranspiration.

Furthermore, the regression deviances, herewith used for intrinsic model validation, decreased from 87 to 50 and from 1 to 0.5 in the case of the re-modelled negative binomial model and the binary model, respectively. This indicated the integrity of the re-modelled regressions (Figures 2B and 3B). The decreases in the deviances and similarly the increases in the significance levels alluded to above (after selecting the three significant variables) essentially confirm the suitability of the selected ecological variables and the importance of variable selection as a precursor for accurate RVF occurrence and risk zone mapping.

3.2. Combining significant variables for risk mapping

Figure 4 shows per pixel distribution maps over the study region for the three variables that were selected for RVF risk mapping; Figure 4A shows 'animal density', 4B illustrates 'small integral', and 4D is 'PC2_ET'. Figure 4C shows the first principal component of ET ('PC1_ET'). The first principal component (Figure 4C) is orthogonal to the second component (Figure 4D), which implies that the two most important principal components often depict divergent landscape features. Herewith, 'PC2_ET' mimics the spatial distribution patterns for 'small integral' somewhat, while there is no apparent visual similarity between

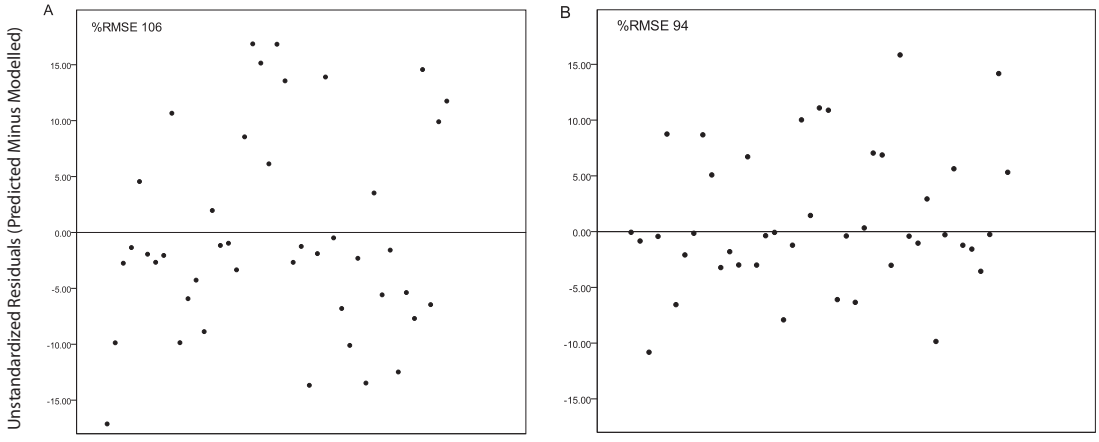

Figure 2. Residual plots for the negative binomial GLM, (A) before and (B) after selecting the three most significant variables and re-running the GLM model. The smaller variance around the zero line after variable selection (B) indicates an improvement in the model and mapping performance.

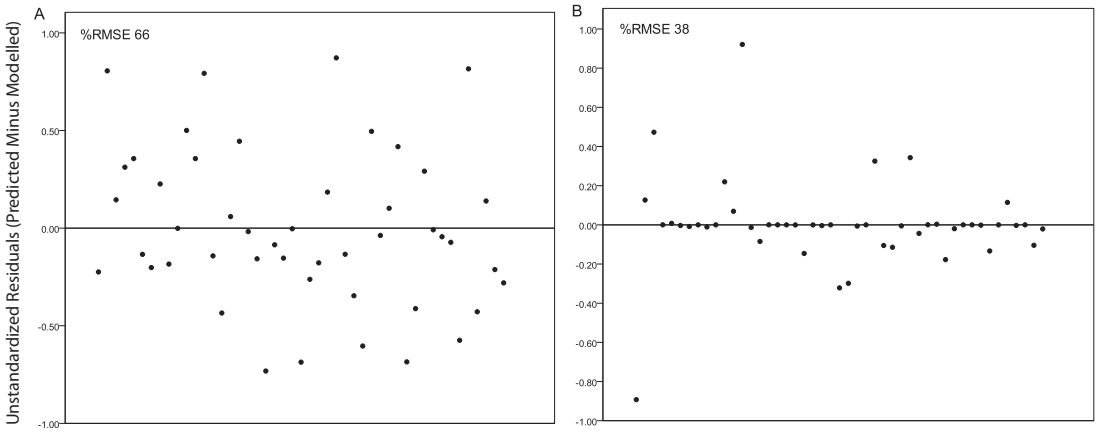

Figure 3. Residual plots for the binary logistic regression GLM, (A) before and (B) after selecting the three most significant variables and re-running the GLM model. The smaller variance around the zero line after variable selection (B) indicates an improvement in the model and mapping performance.

'PC1_ET' and 'small integral' (the second most sensitive variable to RVF occurrence). Higher values for 'PC2_ET', 'small integral', and 'animal density' (illustrated in reddish orange colours in Figure 4) relate to areas of high vulnerability and 'probability' of risk.

Figure 5 shows the risk mapping result derived from amalgamating and weighting the three most significant ecological variables. High risk areas are illustrated in reddish colours, while low risk areas, coloured green, illustrate low RVF occurrence and risk areas. High RVF risk areas were found to be in Tana River, Garissa, Isiolo, and Lamu counties; however, the other counties also showed some small and specific high-risk regions (Figure 5).

Although only 'small integral' was significant in both GLM models (Table 2), the other two variables ('PC2_ET' and 'animal density') were also selected for the risk mapping, since all three variables improved the residual distribution in both GLM models (Figures 2 and 3). Moreover, the results obtained from the binary logistic regression model (in which only 'small integral' was found to be significant; Table 2) can be deemed to be less representative for mapping RVF occurrence over time than the results (variable sensitivity) obtained from the negative binomial GLM model. The negative binomial model results are based on the long-term, i.e. temporally stratified RVF occurrence data.

4. Discussion

The high significance of 'animal density' as a key variable for RVF occurrence suggests that monitoring livestock movement and density, as well as seroprevalence levels in livestock, especially in inter-epidemic periods, may be important in regard to predicting RVF outbreaks.[4]

The statistical importance of the remote sensing-based variables (as found in this study) indicates the potential of remote sensing observations to reduce prediction over-fitting within 'traditional' ecological models.[11] The significance of ET was interesting, since it was hypothesized that ET may be a more meaningful and relevant variable compared to NDVI for mapping the habitat availability of the mosquito vectors. NDVI maps vegetation 'greenness', while ET is sensitive to fluxes from wet soil or flooded areas, as well as chlorophyll active vegetation areas.[24] Mosquito breeding spaces, on a landscape to regional scale, are not

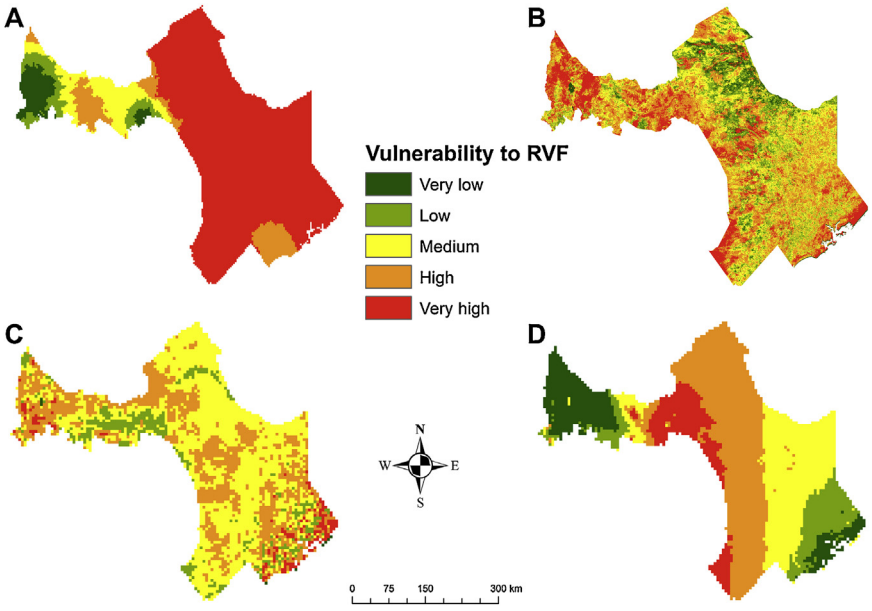

Figure 4. Individual maps showing the three most significant variables (A) 'animal density' (n/km^2), (B) 'small integral', and (D) 'PC2_ET'. (C) Illustrates 'PC1_ET'.

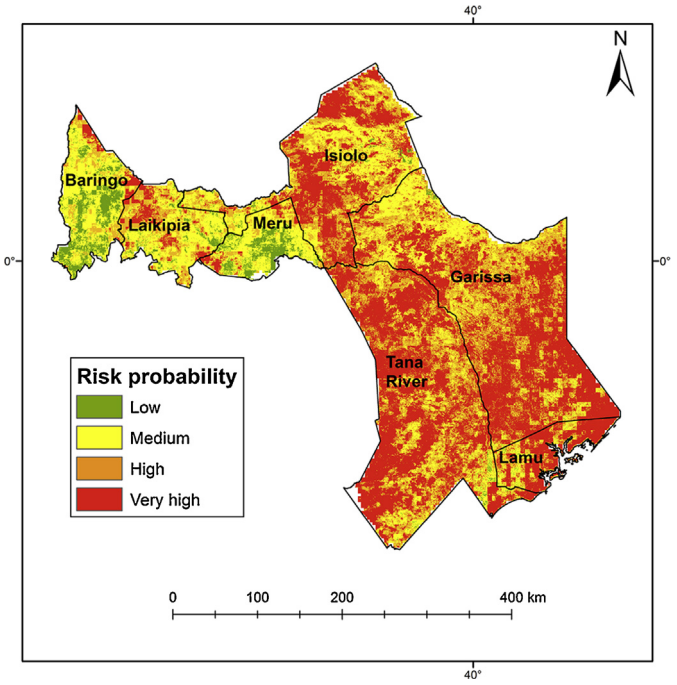

Figure 5. Risk zone map for the study area based on an amalgamation of the variables that were found to be most significant in both GLM models ('animal density', 'small integral', and 'PC2_ET').

only determined by vegetation 'greenness', but also by the actual availability of nearby water and wet soil conditions.[33] The sensitivity of 'small integral' from NDVI was hardly surprising, since the small vegetation integral measures per season increases in NDVI due to rainfall abnormality much more pronounced than 'large integral' or 'season length'. 'Large integral', for instance, exhibits a high baseline NDVI for 'evergreen' areas (i.e., dense woodlands), and an increase in NDVI due to abnormal rainfall would thus not be as apparent.[2,34] This results in a low seasonal difference between the 'greening' response in semi-arid savannas and the 'greening' response in 'other' vegetation zones with generally high chlorophyll activity throughout the year (i.e., dense woodland areas). The suitability of NDVI-based vegetation seasonality (productivity) variables as a trigger mechanism for RVF outbreaks needs to be explored further. In most other RVF modelling studies of eastern Africa,[2,35] the most important factors for RVF occurrence were found to be seasonal flooding and the high abundance of black cotton and alluvial soils, which are known to exhibit high water retention potential.

In this study, the focus was on selecting the most relevant and statistically significant ecological variables as proxies for RVF occurrence and risk zone mapping. Using more spatially explicit, localized, and temporally varying ecological factors from, for instance, remote sensing time-series data streams, the explicitness of RVF occurrence results was improved so that early warning systems with consequent disease intervention efforts can be channelled more precisely.[36]

The spatial RVF risk patterns mapped in this study (Figure 5) were coherent with modelling results attained in other RVF risk mapping studies and for the 2006/2007 RVF outbreak period in particular. Nderitu et al. reported overly high RVF occurrence in Tana River, Garissa, and Ijara using NDVI metrics data and sampling in livestock, mosquitoes, and humans.[16] Britch et al.[10] and Anyamba et al.[9] confirmed that RVF occurrence was highest in the above-mentioned counties during the 2006/2007outbreak. Both studies mentioned Baringo as another RVF 'hot spot' area, while in the present study the county of Baringo, in general, was mapped as a low to moderate RVF risk area, with some isolated high risk areas in between. This underlines the importance of more detailed mapping of risk areas.

To conclude, 'livestock density', 'small vegetation integral', and the second principal component of ET were the most significant determinants of RVF occurrence in Kenya; these were used to map RVF risk zones on a sub-regional scale. However, further studies should also investigate the role of changes in human land use, such as the expansion of irrigated lands, in the propagation of mosquito-borne diseases. This understanding will help to unravel the differences that exist in time and space between enzootic/endemic and non-enzootic/non-endemic areas.

Acknowledgements

We gratefully acknowledge the Swedish International Development Cooperation Agency (SIDA) for support (grant number SWE-2011-016) and the Centre for International Migration and Development (CIM) of the German Development Organization (GIZ). Many thanks also to the numerous colleagues at icipe, Swedish University of Agricultural Sciences, and Umeå University for their technical advice and input.

Ethical approval: The authors confirm that ethical approval was not required for this work since only published disease occurrence data were used and no human or animal sampling was performed.

Conflict of interest: All authors of this manuscript had no competing interests.

References

1. Nguku PM, Sharif S, Mutonga D, Amwayi S, Omolo J, Mohammed O, et al. An investigation of a major outbreak of Rift Valley fever in Kenya: 2006–2007. *Am J Trop Med Hyg* 2010;**83**:5–13.
2. Anyamba A, Chretien JP, Small J, Tucker CJ, Formenty PB, Richardson JH, et al. Prediction of a Rift Valley fever outbreak. *Proc Natl Acad Sci U S A* 2009;**106**:955–9.
3. Drake JM, Hassan AN, Beier JC. A statistical model of Rift Valley fever activity in Egypt. *J Vector Ecol* 2013;**38**:251–9.
4. Lichoti JK, Kihara A, Oriko AA, Okutoyi LA, Wauna JO, Tchouassi DP, et al. Detection of Rift Valley fever virus interepidemic activity in some hotspot areas of Kenya by sentinel animal surveillance, 2009–2012. *Vet Med Int* 2014;**2014**:379010.
5. Hassan OA, Ahlm C, Evander M. A need for one health approach—lessons learned from outbreaks of Rift Valley fever in Saudi Arabia and Sudan. *Infect Ecol Epidemiol* 2014;**4**:20710.
6. Hightower A, Kinkade C, Nguku PM, Anyangu A, Mutonga D, Omolo J, et al. Relationship of climate, geography, and geology to the incidence of Rift Valley fever in Kenya during the 2006–2007 outbreak. *Am J Trop Med Hyg* 2012;**86**:373–80.
7. Muga GO, Onyango-Ouma W, Sang R, Affognon H. Sociocultural and economic dimensions of Rift Valley fever. *Am J Trop Med Hyg* 2015;**92**:730–8.
8. Pin-Diop R, Toure I, Lancelot R, Ndiaye M, Chavernac D. Remote sensing and geographic information systems to predict the density of ruminants, hosts of Rift Valley fever virus in the Sahel. *Vet Ital* 2007;**43**:675–86.
9. Anyamba A, Linthicum KJ, Small J, Britch SC, Pak E, de La Rocque S, et al. Prediction, assessment of the Rift Valley fever activity in East and Southern Africa 2006–2008 and possible vector control strategies. *Am J Trop Med Hyg* 2010;**83**:43–51.
10. Britch SC, Binepal YS, Ruder MG, Kariithi HM, Linthicum KJ, Anyamba A, et al. Rift Valley fever risk map model and seroprevalence in selected wild ungulates and camels from Kenya. *PLoS One* 2013;**8**:e66626.
11. Cord AF, Klein D, Gernandt DS, de la Rosa JA, Dech S. Remote sensing data can improve predictions of species richness by stacked species distribution models: a case study for Mexican pines. *J Biogeogr* 2014;**41**:736–48.
12. Walz Y, Wegmann M, Dech S, Raso G, Utzinger J. Risk profiling of schistosomiasis using remote sensing: approaches, challenges and outlook. *Parasit Vectors* 2015;**8**:163.
13. Murithi R, Munyua P, Ithondeka P, Macharia J, Hightower A, Luman E, et al. Rift Valley fever in Kenya: history of epizootics and identification of vulnerable districts. *Epidemiol Infect* 2011;**139**:372–80.
14. Soti V, Chevalier V, Maura J, Bégué A, Lelong C, Lancelot R, et al. Identifying landscape features associated with Rift Valley fever virus transmission, Ferlo region, Senegal, using very high spatial resolution satellite imagery. *Int J Health Geogr* 2013;**12**:10.
15. McMahon B, Manore C, Hyman J, LaBute M, Fair JM. Coupling vector–host dynamics with weather geography and mitigation measures to model Rift Valley fever in Africa. *Math Model Nat Phenom* 2014;**9**:161–77.

16. Nderitu L, Lee JS, Omolo J, Omulo S, O'Guinn ML, Hightower A, et al. Sequential Rift Valley fever outbreaks in eastern Africa caused by multiple lineages of the virus. *J Infect Dis* 2011;**203**:655–65.
17. Linthicum K, Anyamba A, Tucker C, Kelley P, Myers M, Peters C. Climate and satellite indicators to forecast Rift Valley fever epidemics in Kenya. *Science* 1999;**285**:397–400.
18. Gachohi J, Skilton R, Hansen F, Ngumi P, Kitala P. Epidemiology of East Coast fever (*Theileria parva* infection) in Kenya: past, present and the future. *Parasit Vectors* 2012;**5**:194.
19. Owange NO, Ogara WO, Affognon H, Peter GB, Kasiiti J, Okuthe S, et al. Occurrence of Rift Valley fever in cattle in Ijara district, Kenya. *Prev Vet Med* 2014;**117**:121–8.
20. Sankaran M, Hanan NP, Scholes RJ, Ratnam J, Augustine DJ, Cade BS, et al. Determinants of woody cover in African savannas. *Nature* 2005;**438**:846–9.
21. Munyua P, Murithi RM, Wainwright S, Githinji J, Hightower A, Mutonga D, et al. Rift Valley fever outbreak in livestock in Kenya, 2006–2007. *Am J Trop Med Hyg* 2010;**83**:58–64.
22. Woods CW, Karpati AM, Grein T, McCarthy N, Gaturuku P, Muchiri E, et al. An outbreak of Rift Valley fever in northeastern Kenya, 1997–98. *Emerg Infect Dis* 2002;**8**:138–44.
23. Eastman R, Sangermano F, Ghimire B, Zhu H, Chen H, Neeti N, et al. Seasonal trend analysis of image time series. *Int J Remote Sens* 2009;**30**:2721–6.
24. Mu Q, Zhao M, Running SW. MODIS Global Terrestrial Evapotranspiration (ET) Product (NASA MOD16A2/A3). Algorithm Theoretical Basis Document, Collection 5. NASA; 2013.
25. Dormann CF, Elith J, Bacher S, Buchmann C, Carl G, Carré G, et al. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography* 2013;**36**:27–46.
26. Atzberger C, Eilers PHC. Evaluating the effectiveness of smoothing algorithms in the absence of ground reference measurements. *Int J Remote Sens* 2011;**32**:3689–709.
27. Jönsson P, Eklundh L. TIMESAT—a program for analyzing time-series of satellite sensor data. *Comput Geosci* 2004;**30**:833–45.
28. Robinson T, Thornton P, Franceschini G, Kruska R, Chiozza F, Notenbaert A, et al. Global livestock production systems. Food and Agriculture Organization of the United Nations (FAO); 2011.
29. Plant RE. Spatial data analysis in ecology and agriculture using R. CRC Press; 2012.
30. Cameron AC, Windmeijer FA. An R-squared measure of goodness of fit for some common nonlinear regression models. *J Econom* 1997;**77**:329–42.
31. Williams D. Generalized linear model diagnostics using the deviance and single case deletions. *Applied Statistics* 1987;**36**:181–91.
32. Shahid S, Behrawan H. Drought risk assessment in the western part of Bangladesh. *Natural Hazards* 2008;**46**:391–413.
33. Sang R, Kioko E, Lutomiah J, Warigia M, Ochieng C, O'Guinn M, et al. Rift Valley fever virus epidemic in Kenya, 2006/2007: the entomologic investigations. *Am J Trop Med Hyg* 2010;**83**:28–37.
34. Jonsson P, Eklundh L. Seasonality extraction by function fitting to time-series of satellite sensor data. *IEEE Trans Geosci Remote Sens* 2002;**40**:1824–32.
35. Sindato C, Karimuribo ED, Pfeiffer DU, Mboera LE, Kivaria F, Dautu G, et al. Spatial and temporal pattern of Rift Valley fever outbreaks in Tanzania: 1930 to 2007. *PLoS One* 2014;**9**:e88897.
36. Métras R, Collins LM, White RG, Alonso S, Chevalier V, Thuranira-McKeever C, et al. Rift Valley fever epidemiology, surveillance, and control: what have models contributed? *Vector Borne Zoonotic Dis* 2011;**11**:761–71.